Tianlin Lan. Does the Customer Review of Airbnb Homes Correspond to their Aspect Scores? Sentiment Analysis using a Proposed Sentence-to-aspect Relevance Model. A Master's Paper for the M.S. in I.S degree. April, 2019. 64 pages. Advisor: Stephanie W. Haas

Most Airbnb homes have high ratings with a large proportion of positive reviews from users. However, the Airbnb website only releases the rating score of each aspect for each home. The aspect scores given by each reviewer are not available on the website. It is possible that the overall aspect score does not really reflect users' sentiment as represented in their comments about that aspect.

This paper proposes a methodology for finding the correspondence between aspect scores of Airbnb homes and the sentiments of their reviews. I set the sentiment analysis at the sentence level and proposed a sentence-to-aspect relevance detection approach for subjectivity classification step. The distributions of the sentiment polarities found in aspect-relevant for both the cleanliness and the location aspect show an apparent correspondence between review text and the aspect score.

Headings:

Natural language processing

Unsupervised learning

Sentiment analysis

Opinion mining

Airbnb reviews

DOES THE CUSTOMER REVIEW OF AIRBNB HOMES CORRESPOND TO THEIR ASPECT SCORES? SENTIMENT ANALYSIS USING A PROPOSED SENTENCE-TO-ASPECT RELEVANCE MODEL

by
Tianlin Lan

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

April 2019

Approved by

____

Stephanie W. Haas

TABLE OF CONTENTS

1 INTRODUCTION

The growth of social media and online feedback technologies over the last decade provided customers with more information when making the decision of purchasing products. Online reviews, perform as the electronic word of mouth, have been found to play a significant role. More than 70% of consumers trust reviews in their own online purchasing experience (Bridges & Vásquez, 2018) because of the transparency of reviews. Airbnb, a sharing-economy homestay platform, allows purchasers rate their overall experience, as well as individual aspects such as cleanliness, location, etc. Previous researches pointed out that Airbnb homes have overwhelming positive ratings and reviews from users: nearly 95% of Airbnb homes received an average rating from users of either 4.5 or 5 (maximum) stars (Zervas, Proserpio & Byers, 2015). However, Airbnb website only shows a single score for each aspect to represent all reviewers score. Does the aspect score really represent the all guests' opinion? Is that possible the score covers negative sentiments from guests towards the aspect?

This paper aims to find the correspondence between review text and the scores of the location aspect and the cleanliness aspect of Airbnb homes, based on sentiment analysis. Sentence-level sentiment analysis was deployed in this research with two steps: the subjectivity classification and the sentence sentiment classification. I proposed a sentence-to-aspect relevance detection approach for the first step and used the Stanford CoreNLP[1] in the second step.

The hypothesis of this research is: there exists a correspondence between aspect scores and review text. Airbnb homes with higher aspect scores have larger proportion of positive sentences related to the aspect. The percentage of positive aspect-relevant sentences in reviews increases when the aspect score increases.

[1] https://stanfordnlp.github.io/CoreNLP/sentiment.html

2 LITERATURE REVIEW

This section is the review of the previous relevant literature of this research. The review focuses on two main topics: sentiment analysis, and Airbnb review analysis.

2.1 Sentiment Analysis

Sentiment analysis is about finding out the emotions and sentiments of people towards a wide range of things such as a particular product, the function of a system, an event and so on. The phrase “opinion mining” (Pang & Lee, 2008; Pak & Paroubek, 2010; Esuli & Sebastiani, 2006; Chaovalit & Zhou, 2005) is used interchangeably with “sentiment analysis” in the academic field. The purpose of sentiment analysis and opinion mining is to the classify the sentiment expressed in natural language text. Typically, the classification includes polarity (that is, positive or negative), and may also include strength (very positive, somewhat positive, and so on) (Chaovalit & Zhou, 2005; Liu, 2012).

Research related to sentiment analysis within NLP (natural language processing) field increased rapidly since 2000 (Liu, 2012) when the Web and Internet provided people with techniques and platforms to expressing their opinions through postings, blogs, and comments. The massive amount of opinionated data become valuable assets for different industries to not only listen to the feedback and reactions from their customers or the public, but also make predictions to assist their decision-making process. For example, Bermingham & Smeaton (2011) analyzed the sentiments of Twitter postings to predict people's political attitudes and then predicted the election result based on that analysis. In another project, movie review texts were used in an experiment on sentiment analysis by Joshi, Das, Gimpel & Smith (2010) to predict the opening weekend revenue.

2.1.1 Two Approaches

The studies of sentiment analysis are mainly about two approaches, the corpus-based approach and the lexicon approach (Liu. 2012) which is also known as supervised approach and unsupervised approach respectively (Abdulla, Ahmed, Shehab, & Al-Ayyoub, 2013).

The corpus-based approach requires a considerable volume of data with labels manually annotated by experts. Then, through training and applying machine learning classifiers such as Naïve Bayes (NB) and K-Nearest Neighbor (KNN), researchers have achieved high accuracies in detecting the polarity of a text. Pang, Lee, & Vaithyanathan (2002) implemented various machine learning techniques on movie reviews data from the Internet Movie Database (IMDb) archive of the rec.arts.movies.reviews newsgroup[2]. Their best accuracy was 82.9% when using the "unigrams + bigrams" input features. With the same data, Chaovalit & Zhou (2005) obtained 85.54% accuracy using 3-fold cross validation. However, the corpus-based approach is not very applicable for real-world cases due to the cost of the labeling process. Also, even if the classifier is very effective in a given domain like the movie reviews, the performance varies a lot when

---

[2] http://www.cs.cornell.edu/people/pabo/movie-review-data/

applying the same classifier into other domains (Taboada, Brooke, Tofiloski, Voll & Stede, 2011).

The unsupervised approach, also referred to as the lexicon approach, is more practical. It assigns the polarity value representing positive, negative, neutral to each word according to a prepared lexicon. The lexicon can be created either manually or automatically. The automatic method starts from a set of seed words that contain strong positive or negative associations, such as *excellent* or *abysmal*. The next step is to expand the lexicon by adding adjective words in the whole text that associates with seed words (Taboada, Brooke, Tofiloski, Voll & Stede, 2011). Then, an overall sentiment score of each text or sentence can be computed using polarity values of all words in the text. However, the accuracy may be lower than the supervised approach. Turney (2002) only obtained 65.86% accuracy using the unsupervised approach on 120 movie reviews data from the Epinions[3] website which is less than the accuracy (82.9%) of using supervised approach. Chaovalit & Zhou (2005) also compared two approaches on 100 movie reviews from Movie Vault[4]. They showed that a supervised approach (85.54%) is more accurate than unsupervised one (77%). They suggest that supervised approach costs a tremendous amount of time, while the unsupervised one compromised accuracy a little, but is more practical. Abdulla, Ahmed, Shehab, & Al-Ayyoub (2013) did the same comparison using tweets data in Arabic and reached a similar conclusion. The best accuracy of supervised approaches is 87.2% while the best accuracy of their unsupervised approaches was only 59.6%.

---

[3] http://www.epinions.com/
[4] https://themovievault.net/

Additionally, a hybrid approach merging the two approaches together was also proposed (El-Halees, 2011). El-Halees collected 1143 posts with Arabic statements as documents from education, politics and sports forum. He used the lexicon-based (unsupervised) method to assign the sentiment label for each document, which resulted in 50% accuracy on sentiments of documents. Those labeled documents then were used as training set in the next-step supervised learning model (k-nearest neighbor classifier). The prediction accuracy of predicting sentiments of documents was then increased from 50% to 80%.

2.1.2 Comparing Three Levels of Sentiment Analysis

The sentiment analysis can be performed in different levels, document-level, sentence-level and aspect-level, to offer sentiment information from general to detailed. For each level analysis, this section points out the assumption, summarizes some previous researches and describes its pros and cons.

2.1.2.1 Document-level Sentiment Analysis

Document level sentiment analysis is based on the assumption that each document indicates only one overall sentiment opinion (Pang & Lee, 2008). For example, one review of an Airbnb home only expresses an overall positive or negative attitude towards the home. The challenge of this level of analysis is how to deal with the relations among sentences. If there is one positive sentence and one negative sentence in a document, the coordinating word joining the two sentences may affect the overall sentiment of document. For example, given two sentences, “The service is good. However, it wasn’t worth the luxurious price”, the overall sentiment is a little bit negative. Document-level analysis cannot consider the sentiment transition in the sentence. Studies in this field are still not sufficient to solve this problem, instead, researchers generally neglect this challenge when designing their models. For example, Jiang, Yu, Zhou, Liu & Zhao (2011) considered the semantic relationships within each sentence when proposing and experimenting their document representation based on neural networks. However, they still did not capture the relationships between different sentences. Another distinct disadvantage of document-level sentiment classification is that the only one overall opinion may cover detailed indications in the document.

The disadvantage that document-level sentiment analysis does not deal with fine-grained tasks makes it less suitable for analyzing documents with long length such as newspaper, article, long blogs. However, this level sentiment analysis still works for cases like tweets and reviews that only have a few sentences.

2.1.2.2 Sentence-level Sentiment Analysis

The sentence-level sentiment analysis remedies the defect of document level to some extent. Many applications of this level's analysis come with opinion targets which can be an entity like a product or an aspect of a review, such as the cleanliness of a house or hotel room. By identifying the targets first and then detecting the sentiments of sentences, one could know the opinion polarity (positive or negative) toward the entity or aspect. There are two steps in sentence-level sentiment analysis, the subjectivity classification and the sentence sentiment classification. The first step classifies sentences in a document as being subjective or objective and removes the objective ones based on the assumption that those sentences do not express opinions and sentiments. For example, this sentence in an Airbnb review, "we stayed here for two nights", conveys an objective statement without expressing an opinion about the home. The second step is to do the sentiment classification of each sentence. Each sentence can be regard as a smaller document. However, instead of dealing with the relationships among sentences in document-level, the problem here is about the semantic relationships among words in sentences.

Yu & Hatzivassiloglou (2003) built a Bayesian classifier and then used the sentence similarity for subjectivity classifications. Their model experimented on 2,000 articles from Wall Street Journal (WSJ) from 1987 to 1992. The features they used in the classifier include unigrams, bigrams, the presence of sentiment words, part-of-speech (POS). They measured sentence similarity according to common words, phrases, and WordNet synsets and assumed that subjective (opinioned) sentences are more similar to each other than to the objective sentences. Then, for sentence sentiment classification, they used the unsupervised approach proposed by Turney (2002) to extract phrases with adjectives or adverbs and calculate the sentiment orientation of each phrase using PMI (Pointwise Mutual Information). Their results achieved high precision and recall (F-measure of 91%) on detecting opinion sentences and 90% accuracy in predicting the sentiment polarities. Compared with the original 66% accuracy (Turney, 2002), the improvement is very significant.

Wiebe & Riloff (2005) presented an unsupervised method to develop subjectivity classifiers based on unannotated texts only. In their rule-based subjective classifier, a sentence is classified as subjective if there are two or more strong subjective words. They gathered subjective words from previous published researches, such as: positive and negative n-grams (words and phrases) from research of Dave, Lawrence, Pennock (2003) research, and subjective nouns from the research of Riloff, Wiebe & Wilson (2003). After experimenting on 535 texts from publications of the Foreign Broadcast Information Service (FBIS), their model achieved 73.1% subjectivity precision and 66.2% subjectivity recall. Barbosa & Feng (2010) proposed an approach in the subjectivity classification of postings on Twitter that included special features of Twitter, retweets, hashtags, smileys, etc. Jiang, Yu, Zhou, Liu & Zhao (2011) improved the accuracy of sentiment classification for Twitter texts from 78.8% to 85.6% by adding target-dependent features for sentences in their research.

A sentiment lexicon with words indicating sentiment polarities is widely used in sentence sentiment classification. Hu & Liu (2004) proposed a lexicon-based algorithm of aspect-based sentiment analysis to detect the sentiment polarities of sentences. They created the sentiment lexicon through WordNet using given sentiment words and their synonyms and antonyms. Then, for every word in a sentence, they assigned polarity, 1 or -1 (representing positive or negative). The sentence sentiment then can be computed as the sum of scores of all words in it. Kim & Hovy (2004) used the same method as Hu & Liu (2004) in their research with a small adjustment: instead of calculating the sum of word scores in a sentence, they multiplied the scores of sentiment words in sentence. Ding, Liu & Yu (2008) proposed a technique to determine the sentiment polarity of an aspect in a review sentence by considering the distance (the number of words) between opinion words and the target word. Opinion words were obtained through a bootstrapping process of WordNet. The target word means the aspect or feature the sentence talks about such as “battery” in the sentence, “The battery of this camera lasts very long”. The polarity score of the opinion target was computed using the following formula.

$$\text{score(aspect)} = \sum_{opinion\ words\ in\ sentence} \frac{sentiment\ value\ of\ opinion\ word}{distance(opinion\ word, aspect)}$$

They used their method on customer reviews of 8 digital products and achieved 92% precision, 91% recall and 0.91 F-score. The results are better than the results of using methods proposed by Hu & Liu (2004), of 93% in precision, 76% recall and 0.83 F-score.

2.1.2.3 Aspect-level Sentiment Analysis

Aspect-based sentiment analysis (ABSA) extracts opinions more detailed than the other two levels by decomposing the document or sentence into aspect-relevant corpus of words or phrases. The assumption is in accord with the cases in reality that each document or even each sentence may express opinions for multiple targets/aspects. For example, a sentence in the Airbnb review could not only talk about the cleanliness aspect but also the accuracy aspect and the location aspect. The ABSA extracts the mentions of aspects from documents and uses phrases in the documents that have dependency with extracted aspects to do sentiment classification. Hu and Liu (2004) assumed that the nouns represent aspects and those nouns repeated in different reviews are more important aspects. They extracted nouns and noun phrases with higher occurrence frequencies as aspects. The extracted aspects were then tagged with the closest adjectives. Those adjectives were defined as the opinion words in the research. By determining the sentiment of opinion words, they predicted the sentiments of aspects. Other studies improved the aspect extraction method. Researchers (Blair-Goldensohn, Sasha, Hannan, McDonald, Neylon, Reis & Reynar, 2008) used the noun frequency method on reviews of local services and optimized the method by adding dynamic aspect extractor in the aspect extraction step. Dynamic aspect extractor can extract infrequent aspects. Through combining both frequent and infrequent aspects in the experiment, they increased the F-score of both restaurant (68.9 to 77.4) and hotel (65.9 to 74.8) data. Another approach used to improve Hu and Liu's method is to improve the extraction of sentiment words for aspects through the dependency relation. Instead of using the closest adjectives to tag aspects, Kim & Hovy (2006) proposed a semantic role labeling framework to construct dependencies between aspect word (named as target word in their research) and other words and phrases. They labeled semantic roles including phrase type, head word, parse tree path, position and voice for words in a sentence. Their experiment data were 8,256 and 11,877 sentences that are associated to opinion bearing frames for verbs and adjectives from FrameNet[5] annotation data. Using the semantic role labeling method, their system improved F-score for both adjective (38.2% to 70.3%) and verb target words (30.4% to 66.5%).

From the probability perspective in considering the aspect extraction, researchers used topic modeling in ABSA. Topic model discovers the abstract "topics" that occur in a collection of documents. In sentiment analysis, the "topics" can be regarded as aspects. Lu, Ott, Cardie & Tsou (2011) compared the performances of LDA (Latent Dirichlet Allocation) model, Local LDA model, Multi-Grained LDA model and Segmented Topic Models (STM) in sentiment analysis. The data they used for their aspect-labeling experiment were 73,495 restaurant reviews with the aspect ratings on food, service and ambiance aspects. They suggest that those topic models have good performance given that all models lead to higher than 0.7 F-score for all three aspects. However, only the

[5] https://framenet.icsi.berkeley.edu/fndrupal/fulltextIndex

best accuracy (80.3%) by using Local LDA is comparable to the result (83.0%) of using supervised SVM classifier while other models only have accuracies: 47.7% (LDA), 76% (MG-LDA) and 79.4% (STM).

2.1.2.4 Comparison

Compared with document-level analysis, sentence-level sentiment analysis focuses more directly on the opinion targets. Sentence-level sentiment analysis removes the unopinionated sentences in a document so that reduces the noise of those sentences and thus increases the performance of sentiment detection. However, it generally requires additional information, the opinion target, to do the subjectivity classification step. Only knowing the sentiment expressed by a sentence is useless if which target the sentence talks about is unknown. So, sentence-level analysis is more feasible in cases like online reviews, tweets which have specific opinion targets rather than texts in forums, blogs or newspapers. Aspect-level sentiment analysis performs better when encountering complex sentences that express multiple aspects with different sentiments. It orients phrases or words in sentences to their relevant aspects. So, even if a sentence contains multiple aspects, one can use the relevant text of each aspect to determine its polarity. However, because the ABSA method extracts aspects directly from texts, if one wants to analyze the texts for a given aspect, it is possible that the given aspect is not included in the extracted aspects.

2.2 Airbnb Studies

Zervas et al. (2015) analyzed the distribution of score ratings of Airbnb homes and compared it with rating distributions of traditional accommodation industry (hotels) and another sharing economy platform, TripAdvisor. They found the proportion of positive ratings on Airbnb is overwhelmingly more than the other two and suggest the underlying reasons from the qualitative perspective. They suggest some behaviors of hosts of Airbnb may cause the too-positive phenomenon: rejecting customers that they identify as unsuitable on their own, preempting negative reviews and creating new property page to reset previous negative scores. Bridges & Vásquez (2018) analyzed 400 reviews of Airbnb from both guest and hosts (each with 200 reviews). They manually coded and categorized those reviews as either positive or negative. By coding the positive and negative corpus of reviews, they got the result that 93% of Airbnb reviews were positive which suggests a positivity bias in Airbnb reviews. They also supported one of the conclusions from Zervas et al. (2015) that guests maybe unwilling to write negative comments about the hosts. Their results find that negative reviews on Airbnb are more about the property or the location rather than the host, which suggests reviewers tend to avoid negative comments of host in their reviews.

Cheng & Jin (2019) implemented aspect-based sentiment analysis on Sydney Airbnb review data using Leximancer (a text mining software[6]) and compared the extracted aspects with concepts gathered from previous hotel stays literatures. Their results suggest that when assessing the experience, guests of Airbnb use basically the same attributes used by guests reviewing hotels, but the priority of attributes is different. Location, amenities, and host are aspects that Airbnb guests mostly care about. They also suggest that there exists a positivity bias in Airbnb reviews.

---

[6] http://info.leximancer.com/

Hoffen, Hagge, Betzing, & Chasin (2017) used two sets of data: Airbnb official review data and comments from Twitter with hashtag “Airbnb”. They extracted the sentiment polarities of aspects using Stanford coreNLP model first. Then they compared the occurrence frequencies of aspects with different sentiment polarities in both data corpuses. Their results suggest that the occurrences of positive aspects in both datasets take largest proportion and negative aspects take the least. However, the proportion of positive aspects in Airbnb official dataset is 16.54% larger than in Twitter dataset while the proportion of negative aspects is 8.71% less. This suggests a greater positive bias of reviews on Airbnb website than other review sources.

3 METHODOLOGY

In this section, I describe the process of conducting the aspect-based sentiment analysis in this research. The following subsections:

- describe the dataset used in this research,
- elaborate the data cleaning and selection process,
- propose and validate the sentence-to-aspect relevance detection model, and
- explain the method of getting the sentiment polarities values and the analysis method of sentiments.

The goal is to observe the correspondence between customers' review texts and the numeric aspect scores of Airbnb homes. All python codes have been presented in a GitHub repository[7].

3.1 Data

3.1.1 Initial Dataset: The Officially-Released Data of Seattle by Airbnb

The initial dataset of this research includes two csv tables, listing.csv and reviews.csv, which contains the homes and reviews information in Seattle city released by Airbnb official account on Kaggle[8]. It can be assumed that the dataset is authentic and reliable given that it was released by Airbnb. The dataset also represents contemporary

---

[7] https://github.com/Lynnlan/Airbnb-review-analysis-Correspondence-between-Aspect-scores-and-Review-text
[8] https://www.kaggle.com/airbnb/seattle

Airbnb data given the scraped date of 01/04/2016. Table 1 shows the detailed contents of the two tables.

Attributes / Data Table	Size (rows × columns)	Scraped Date	Columns
Listings.csv	3,818 × 92	01/04/2016	[listing_id, name, review_scores_rating, review_scores_accuracy, review_scores_checkin, review_scores_cleanliness, review_scores_location, review_scores_value, review_scores_communication, … (other 83 columns)]
Reviews.csv	84,849 × 6	06/07/2009 to 01/03/2016	[review_id, listing_id, reviewer_id, reviewer_name, date, comments]

Table 1: Basic Information of Dataset

In summary, listings.csv provides the home description information including the distinct listing_id, price, accommodation, etc. and the score information of the homes such as overall score rating, and 6 aspect scores. The reviews.csv includes the distinct ids of each review, the listing_id reviews were written for, reviewer information and comments text. Two tables can be interlinked by the “listing_id”. One Airbnb can have multiple reviews while each review is only about one home.

3.1.2 Data Cleaning of the Initial Dataset

This section is about the regular data cleaning process including removing empty data, and selecting necessary columns from the dataset for this research.

I removed 660 homes from the 3818 homes included in listings.csv, the table of homes information, because they were missing aspect scores. Next, I extracted the columns needed for this research. I focus on the location aspect and the cleanliness aspect, so out of the 92 columns in this table, I used the following 4 columns:

- listing_id: the unique id of each home
- number_of_reviews: the number of reviews for each home
- review_scores_cleanliness: the average cleanliness score for each home
- review_scores_location: the average cleanliness score for each home

As for the reviews.csv, the table of reviews information, I removed 1119 reviews from 84849 reviews. The category of removed reviews are:

- Reviews automatically posted by Airbnb system with the same text: "The host canceled this reservation the day before arrival. This is an automated posting." (57 reviews)
- Reviews that are not in English, for example, "위치가 너무 좋았습니다. 스티브는 너무 친절했습니다.\r\n 가격도 저렴합니다.\r\n 다시 시애틀에 간다면 또 이용하겠습니다". (1032 reviews)
- Reviews with meaningless text such as only punctuation or a smiley. For example: ".", ":)" etc. (30 reviews)

Out of the 6 columns in this table, I used the following 3 columns:

- listing_id: the unique id of each home
- review_id: the unique id for each review
- comments: the text of reviews

Then, I merged two cleaned tables together using the "listing_id" column. After cleaning and merging, in the end, the data set contains 3,155 Airbnb homes and 83,682 reviews.

3.1.3 Data Selection for the Analysis of Two Aspects

This section describes the sampling process to select the homes and reviews for the sentiment analysis of the location and cleanliness aspects. The goal was to include a sufficient and comparable number of Airbnb homes and reviews for each score for further sentiment analysis. Some scores were grouped together and analyzed as a single group when the numbers of records were too small to be analyzed separately.

3.1.3.1 The Subset of Data for Cleanliness Aspect

Table 2 shows the uneven distribution of number of homes and reviews by different scores of cleanliness aspect. The distribution of number of records for different cleanliness scores is highly skewed. The number of records of cleanliness scores from 3 to 7 is far less than the number of records of higher cleanliness scores, 9 and 10. To solve this, I firstly combined homes with cleanliness score 7 or lower into a single group "0-7" and kept homes with 8, 9, 10 scores as independent groups. This provided more balanced groups, allowing me to make comparison of sentiment polarities distribution results. This should be a reasonable way because the previous findings suggest that people have a positive bias when commenting reviews on Airbnb. If it is indeed the case, the data of "0-7" group possibly represents negative opinion from guests.

Scores of Cleanliness	Number of Homes	Number of Reviews
3.0	1	2
4.0	4	9
5.0	5	23
6.0	30	107
7.0	40	511
8.0	183	3,569
9.0	740	22,197
10.0	2,152	57,262

Table 2: Distribution of Homes and Reviews by Cleanliness Score

Then, for the large number of records of cleanliness scores higher than 7, I created a random sample of 75 homes for each score group to reach approximately 2000 (the acceptable range is 1900 to 2100) number of reviews for further sentiment analysis. For the purposes of this study, I assumed that 2000 reviews would be a large enough sample for analysis. Given the average number of reviews of each home is 26.53, approximately 75 homes were needed for each group by cleanliness score to get 2000 reviews. 75 is also a reasonable number of homes that can be expected to show the sentiment distribution patterns within a score. As for the “0-7 score” group, considering the small data size, I chose to keep all the records. Table 3 shows the detailed information of subset in this section:

Scores	0 – 7	8	9	10
Selected number of homes	28	75	75	75
Number of reviews of selected homes	652	2026	2014	1984

Table 3: The Information of Subset Data for Cleanliness Aspect

3.1.3.2 The Subset of Data for Location Aspect

Table 4 shows the distribution of number of homes and reviews by different scores of location aspect. We can see it is similar to the distribution in cleanliness aspect. There is an huge inequality of the number of records among lower location scores (7 or lower) and higher location scores (9, 10). The number of records of location score 8 are appropriate.

Scores of Location	Number of Homes	Number of Reviews
4.0	1	1
6.0	8	13
7.0	19	150
8.0	128	1,833
9.0	884	29,151
10.0	2,115	52,534

Table 4: Distribution of Homes and Reviews by Location Score

So, I followed the same procedure of selecting subset of data in cleanliness aspect. The only difference with cleanliness data selection was the data in location score 8 group. There are only 1833 number of reviews which is less than 2000, I did not make random selection here but instead kept all the records. Table 5 shows the detailed information of subset in this section.

Scores	0 – 7	8	9	10
Selected number of homes	28	128	75	75
Number of reviews of selected homes	164	1833	2017	1967

Table 5: The Information of Subset Data for Location Aspect

3.2 Sentence-to-Aspect Relevance Detection Approach

3.2.1 The Relevance Lexicon of Aspects

The lexicon of aspects is a bag of words related to each aspect which will be used to detect whether a sentence is relevant to the given aspect. I used the WordNet library in Python, which offers a large lexical database of English containing 117,000 synsets[9] of

[9] https://wordnet.princeton.edu/

words. Each synset covers a set of synonyms that share a specific and distinct concept by similar semantic definition. All synsets are linked with each other by semantic relationships such as hyponymy, antonymy, and so on. In addition, by reading through the given definitions of synsets of one word, one can deal with the polysemous problem by only selecting the useful synsets and ignoring others. For example, the synset “goodly.s.01” of word “tidy” with the meaning “large in amount or extent or degree” is not the meaning I expected to use in cleanliness aspect analysis. After reading the definition, I did not include the words in this synset in my relevance lexicon.

The first step of extracting relevance lexicon of a certain aspect is to determine the keywords of this aspect as the starting point of using WordNet. In this research, I extracted keywords of each aspect from the “how do star ratings work” webpage of Airbnb which gives the brief official description. For example, the description of the aspect location is “How did guests feel about your neighborhood?[10]”; therefore, I used “location” and “neighborhood” as the keywords of location aspect.

The next step is finding the bag-of-words related to the given aspect using its keywords as the starting point in the WordNet’s synsets. In this research, there are 3 steps:

a) Filtering the synsets of all keywords. Filtering the synsets of keywords is necessary because there are some synsets of the keywords do not mean the concept of the word in the Airbnb context. For example, as shown in Figure 1, the location.n.04,

---

[10] https://www.airbnb.com/help/article/1257/how-do-star-ratings-work

localization.n.01, placement.n.03 are not necessarily in accord with the Airbnb's definition of aspect location. So, synonyms in those synsets will be ignored.

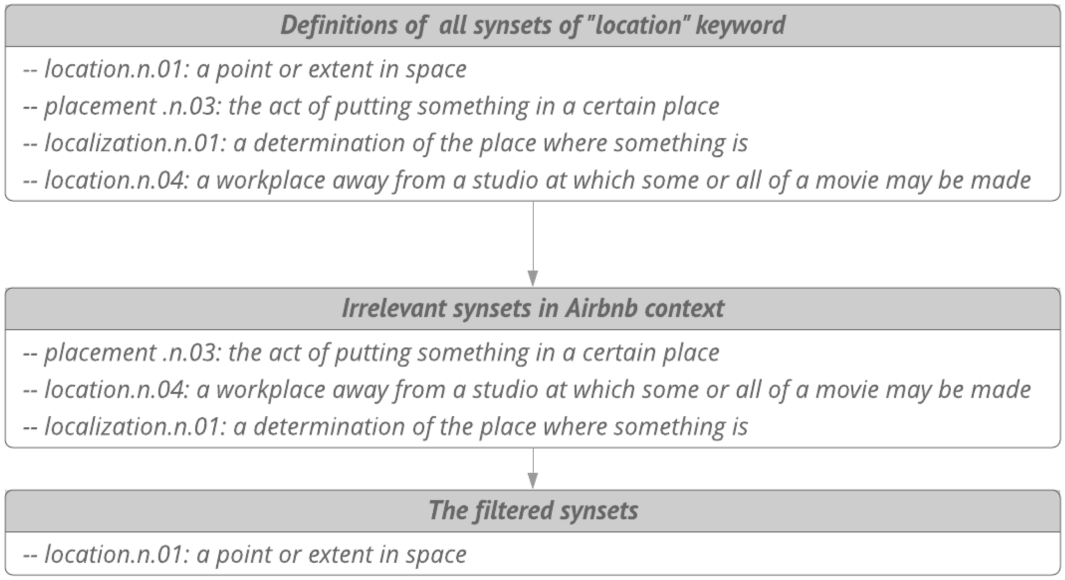

Figure 1: The Synsets Selection Process of Keyword “Location”

b) Retrieve both synonyms and antonyms of the filtered synsets in bag-of-words. When people comment about one aspect, they may describe their opinions by expressing the opposite evidence. For example, comments like “The apartment was clean and quiet” is relevant to the cleanliness aspect because it directly mentions the word “clean”. Comments like “The bathroom was very dirty with dust and hair in the tub and on the floor” also is relevant even though it doesn’t contain the symptoms of “cleanliness”.

c) Investigate words in the broader or narrower semantic level of words in step b. I went through the process of extracting extra synonyms and antonyms of the retrieved words in step b aiming to enlarge the relevance lexicon of aspect. However, after I proofread the additional words, I found a large proportion of the broader or narrower terms were not actually relevant to the given aspect. See the example in the

Figure 2: The added words from broader and narrower semantic levels do not necessarily share the conceptual meaning of the "cleanliness" aspect in Airbnb.

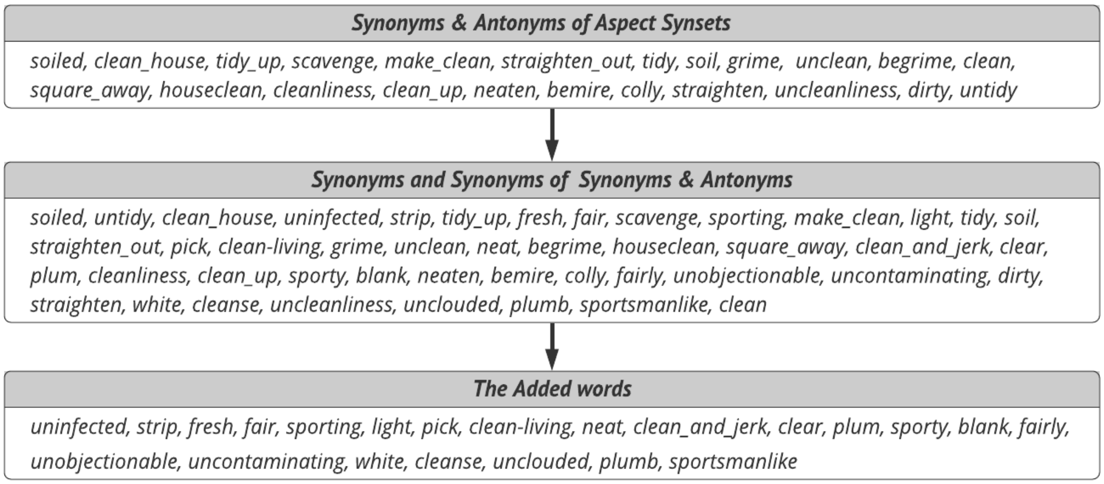

Figure 2: Words in Broader and Narrower Semantic Levels of "cleanliness"

Table 6 shows the whole process and results of implementing above steps in creating the relevance lexicon of aspects location and cleanliness.

Aspect	Keywords	Synsets	Filtered Synsets	Synonyms in Filtered Synsets	Antonyms	Relevance Lexicon
Location	Location, neighborhood	location.n.01, placement.n.03, localization.n.01, location.n.04, vicinity.n.01, neighborhood.n.02, region.n.04, neighborhood.n.04	location.n.01, vicinity.n.01, region.n.04, neighborhood.n.04	locality, location, neighborhood, neighbourhood, region, vicinity, neck_of_the_woods	None	locality, location, neighborhood, neighbourhood, region, vicinity, neck_of_the_woods
Cleanliness	clean, tidy, cleanliness	cleanliness.n.01, cleanliness.n.02, tidy.n.01, tidy.v.01, tidy.a.01, kempt.s.01, goodly.s.01, clean_and_jerk.n.01, clean.v.01, clean.v.02, houseclean.v.01, cleanse.v.01, clean.v.05, clean.v.06, clean.v.07, clean.v.08, scavenge.v.04, clean.v.10, clean.a.01, clean.s.02, clean.s.03, clean.s.04, clean.s.05, clean.a.06, clean.a.07, clean.a.08, uninfected.s.01, clean.s.10, clean.s.11, blank.s.01, clean.s.13, clean.s.14, clean.s.15, clean.s.16, clean.s.17, clean.s.18, clean.r.01, fairly.r.03	clean.v.01, houseclean.v.01, clean.v.05, clean.v.08, scavenge.v.04, clean.a.01, cleanliness.n.01, cleanliness.n.02, tidy.v.01, tidy.a.01	houseclean, tidy_up, scavenge, clean_house, make_clean, straighten_out, cleanliness, square_away, clean, neaten, clean_up, straighten, tidy	begrime, bemire, colly, dirty, grime, soil, soiled, unclean, uncleanliness, untidy	soiled, clean_house, tidy_up, scavenge, make_clean, straighten_out, tidy, soil, grime, unclean, begrime, square_away, houseclean, cleanliness, clean_up, neaten, bemire, colly, straighten, uncleanliness, dirty, untidy, clean

Table 6: The Process and Results of Creating Relevance Lexicon for Aspect Location and Cleanliness

3.2.2 Detecting the Relevance between Sentences and Aspects

For each aspect, I used the sentence-aspect relevance detection model described below to create the corpus of all relevant sentences. The corpus of sentences of each aspect were then used for the sentiment analysis and sentiment distribution. In the sentence-aspect relevance detection model, the input variables are one sentence and the relevance lexicon of aspects. One parameter, the relevance score, will be used to adjust the model. Two outputs, word-similarity-score and sentence-aspect-similarity will be computed during the process.

Firstly, for each word in each sentence, I computed the “word-similarity-score” towards the keywords of one aspect. The value of word-similarity-score ranges from 0 to 1, with 0 representing poor similarity and 1 representing high similarity between two words. For example, relevant words in the sentence, “The bathroom was not clean unfortunately”, result in the similarity scores shown in Table 7.

Words in sentences / Aspect Lexicon	bathroom	not	clean	…	unfortunately
tidy	0.39977372	0.24361573	0.61905617	…	0.3180887
…	score	score	score	…	score
dirty	0.42091253	0.38980597	0.58285207	…	0.3533205

Table 7: Example of word-similarity-score in a Sentence

I used the similarity method in spaCy, an open-source NLP library to compute it. I chose to load the “en_core_web_md” model from the 4 models in spaCy for English language processing. The chosen model contains “English multi-task CNN trained on

OntoNotes, with GloVe vectors trained on Common Crawl[11]" and "685k keys, 20k unique vectors (300 dimensions)". It was also suggested by its author to use for written text (blogs, news, comments). So, this model can be considered naturally suitable to analyzing online Airbnb review texts.

Then, I defined a "threshold relevance score" to control sensitivity of this sentence-aspect relevance model: a sentence is recognized as aspect relevant when there is at least one word has the word-similarity-score larger than the "threshold relevance score" parameter.

3.2.3 Threshold Relevance Score Selection

I selected several Airbnb homes in the dataset with more than 50 reviews each and did manual validation for them using different relevance scores 0.5, 0.6 and 0.7. For example, for home 30712 has 64 reviews which consists of 452 sentences, I used the proposed model to find sentences relevant to the cleanliness aspect. I tried 3 different threshold relevance score parameters: 0.5, 0.6 and 0.7 resulting in 3 corresponding corpuses of relevant sentences. Every sentence in those 3 corpora contains at least 1 word that has the word-similarity-score larger than threshold relevance score. As shown in Table 8, the model detected 221, 43, and 31 cleanliness-relevant sentences for the three threshold relevance scores, respectively. For each sentence in each corpus, I manually checked the accuracy of results detected by model and summarized the precision.

---

[11] https://spacy.io/models/en

Threshold Relevance Score	Number of Sentences Detected as Relevant	Number of Relevant Sentences Checked Manually	Precision
0.5	221	47	21.26%
0.6	43	43	100%
0.7	31	31	100%

Table 8: Precision of Relevant Sentence Corpuses by Running Sentence-Aspect Relevance Model with Different Relevance Score on Home30712

I found that the 0.5 corpus didn't perform well with the accuracy 21.26%. The reason may be that the word "clean" itself is a positive adjective which in the spaCy similarity model is recognized to have word-similarity-score in range 0.5 – 0.6 with other positive words like "wonderful", "nice", "great", etc. The outcome of this is that sentences like the followings were all detected as relevant sentences of "cleanliness":

- "Overall, the neighborhood is nice, Al, and his daughter, are very welcoming and Seattle is always great."
- "The owner is super nice and accommodating."
- "I had the whole place to myself so that was really nice."

When I set the relevance score as 0.6, performance improved to 100% precision based on my proofreading for this home 30712. However, when use 0.7 as relevance score, although the results' precision is also 100%, the number of sentences in the corpus dropped a little. The underlying problem here was the balance between precision and recall. Using the relevance score 0.5 definitely guaranteed the recall of the result however compromise the precision while 0.7 by contrast has pretty high precision but impaired the recall. Therefore, I used 0.6 in the sentence-aspect relevance model this research.

3.3 Sentiment Analysis of Sentences

3.3.1 Determine the Sentiment Polarity of Each Sentence

This section is to find out the sentiment polarities of aspect-relevant sentences identified as being relevant to an aspect. I chose to use Stanford CoreNLP which has the underlying technology based on a new type of Recursive Neural Network (RNN) that builds on top of grammatical structures. The model provided the sentiment polarities values for input sentences. The values 4, 3, 2, 1, 0 represent the sentiment polarities, very positive, positive, neutral, negative and very negative respectively. Compared with other widely-used sentence sentiment analysis tools, such as the NLTK Text-Processing API which uses the lexicon approach, Stanford CoreNLP's algorithm is more advanced and comprehensive because its learning model actually builds up a representation of whole sentences based on the sentence structure. It computes the sentiment based on the sentence structure rather than simply summing up the positive or negative point of each word.

3.3.2 Analysis of the Sentiments of Aspect-Relevance Sentences

For each home, I counted the number of sentences per sentiment polarity, for example, home30712 has 24 sentences with sentiment polarities negative, 9 sentences with neutral sentiment, and 9 with positive sentiment, in its cleanliness aspect-relevant corpus of sentences. All statistics are in Appendix I. The hypothesis here is that the numeric aspect scores of homes represent customers' opinions towards homes. Intuitively, homes with higher aspect score should receive larger percentage of positive comment sentences. I analyzed these statistics from the following perspectives:

- Comparing the sentiment polarities distribution of sentences among score groups in each aspect. The assumption is that as the score increases, the percentages of sentences with very negative and negative sentiment polarities decrease and the percentages of positive and very positive increase. I calculated two variables for analysis:
    - The percentages of sentences for each sentiment polarity value. This shows the change of the percentages across score 0 – 7, 8, 9 and 10.
    - The average percentages of the sentiment polarities percentages of homes with the same aspect score. This shows the change of the average percentages across score 0 – 7, 8, 9 and 10. The following formula shows how to calculate this variable.

$$\text{AvgPct(polarity)} = \frac{\sum_{homes} home's\ polarity\ percentage}{number\ of\ homes}$$

- Observing the sentiment distribution for homes with the same aspect score. The assumption here is that homes with the same score have similar percentage distribution of different sentiment polarities (positive, neutral, negative, etc.) in review texts. For homes with the same aspect score, such as all homes with cleanliness score 8, I want to observe the distribution pattern of sentiment polarities. I calculated the percentages of each sentiment polarities for each home and plot it in stacked bar chart. Each stacked bar represents one home and the stacks in it represents the percentages of number of sentences with different sentiment polarities.

4 RESULTS AND DISCUSSION

Results are presented in this section. First, I report characteristics of the data subsets for the location and cleanliness aspects. Next, I report results of the sentiment analysis of sentences associated with each aspect. Finally, I discuss the tendency and patterns of sentiment polarities distribution.

4.1 Statistics of Data Subsets for Aspects

I choose data subsets for both location and cleanliness aspects and used the sentence-aspect relevance detection model for reviews in subsets to get a corpus of aspect relevant sentences of the given aspect. Table 9 shows the statistics of all steps for the 2 data sets. We can see that the data sizes of both aspects are pretty close in the “#homes”, “#reviews”, “#sentences” and “#sentences per review”. However, there is a difference in the number of aspect-relevant sentences (“#relevant sentences”) where location occurs approximately twice as frequently as cleanliness. This suggests that people talk more about the location aspect than the cleanliness aspect when writing reviews for Airbnb homes.

ASPECT	CLEANLINESS	LOCATION
# HOMES	305	306
# REVIEWS	6676	5981
# SENTENCES	35575	32996
# SENTENCES PER REVIEW	5.33	5.51
# RELEVANT SENTENCES	2518	5335

Table 9: Statistics of Data Subsets for Cleanliness and Location Aspect

Table 10 shows more detailed information of both aspects' data grouped by scores. For both aspects, there exist some homes without aspect-relevance sentences. However, the number is acceptable because after removing those, there remains a sufficient number of homes for analysis. An observation is that guests tend to mention the aspect when writing reviews if they rate the aspect a higher score. As the table shows, 21.15% and 26.67% of homes with lower scores on the cleanliness aspect (0-7 or 8) do not have aspect-relevant sentences. In contrast, only 9.3% and 13.3% of homes with scores of 9 or 10 are missing relevant sentences. The same pattern is seen for the location aspect, 14.28% and 16.4% of homes with lower scores on the location aspect (0-7 or 8) score do not have aspect-relevant sentences. In contrast, only 4% of homes with the location aspect score of 9 are missing relevant sentences. All homes with score of 10 have sentences in their reviews that are relevant to the location aspect.

ASPECT	STATISTICS \ SCORE	0 – 7	8	9	10
CLEANLINESS	Total # homes	80	75	75	75
	Homes without aspect-relevant sentences # (%)	27 (21.25%)	20 (26.67%)	7 (9.33%)	10 (13.33%)
	Homes with aspect-relevant sentences # (%)	53 (66.25%)	55 (73.33%)	68 (90.67%)	65 (86.67%)
	Total # aspect-relevant sentences	276	762	696	784
	Avg # relevant sentences per home	5.21	13.85	10.23	12.06
LOCATION	Total # homes	28	128	75	75
	Homes without aspect-relevant sentences # (%)	4 (14.28%)	21 (16.41%)	3 (4%)	0 (0%)
	Homes with aspect-relevant sentences # (%)	24 (85.72%)	107 (83.59%)	72 (96%)	75 (100%)
	Total # aspect-relevant sentences	183	1514	1739	1899
	Avg # relevant sentences per home	7.63	14.15	24.15	25.32

Table 10: Statistics of Data Subsets for Cleanliness and Location Aspects by Rating Scores

Another way of looking this is that for homes whose reviews do contain relevant sentences, ones with higher scores have a larger average number of relevant sentences than those with lower scores. As for the cleanliness aspect, each home with lower aspect score (0-7) only have an average of 5.21 relevant sentences, while other homes with higher aspect scores (8, 9, 10) have more than 10 relevant sentences on average. Similarly, for the location aspect, homes with lower aspect scores (0-7) only have 7.63 relevant sentences on average, while other homes have more. Homes with aspect score of 8 have 14.15 aspect relevant sentences on average, homes with 9 or 10 aspect scores have more than 24 aspect relevant sentences on average.

4.2 Sentiment Statistics for Aspects

4.2.1 Overall Sentiment Polarities Distribution of Aspects

After extracting sentiment polarities of all sentences, I calculated the number of sentences and the percentages of all polarities as shown in Table 11.

ASPECT	**POLARITIES**	**# SENTENCES (%)**			
		Score 0 – 7	Score 8	Score 9	Score 10
CLEANLINESS	very negative	9 (3.26%)	9 (1.18%)	5 (0.72%)	1 (0.13%)
	negative	154 (55.80%)	308 (40.42%)	167 (23.99%)	138 (17.60%)
	neutral	37 (13.41%)	101 (13.25%)	39 (5.60%)	42 (5.36%)
	positive	68 (24.64%)	311 (40.82%)	411 (59.06%)	511 (65.18%)
	very positive	8 (2.90%)	33 (4.33%)	74 (10.63%)	92 (11.73%)
LOCATION	very negative	4 (2.19%)	8 (0.53%)	9 (0.52%)	7 (0.37%)
	negative	78 (42.62%)	547 (36.13%)	524 (30.13%)	481 (25.33%)
	neutral	23 (12.57%)	199 (13.14%)	172 (9.89%)	159 (8.37%)
	positive	75 (40.98%)	694 (45.84%)	924 (53.13%)	1067 (56.19%)
	very positive	3 (1.64%)	66 (4.36%)	110 (6.33%)	185 (9.74%)

Table 11: Sentiment Polarities Distribution Statistics of Cleanliness and Location Aspect grouped by Aspect Scores

These statistics show that there is tendency of sentiment polarities of review texts to correspond to aspect scores. As the aspect score increases, the percentages of very negative, negative and neutral decrease. Similarly, the percentages of very positive, positive and increase. Figures 3 and 4 show the relationship between rating and sentiment polarity. The tendency is strongest for the positive and negative sentiment categories.

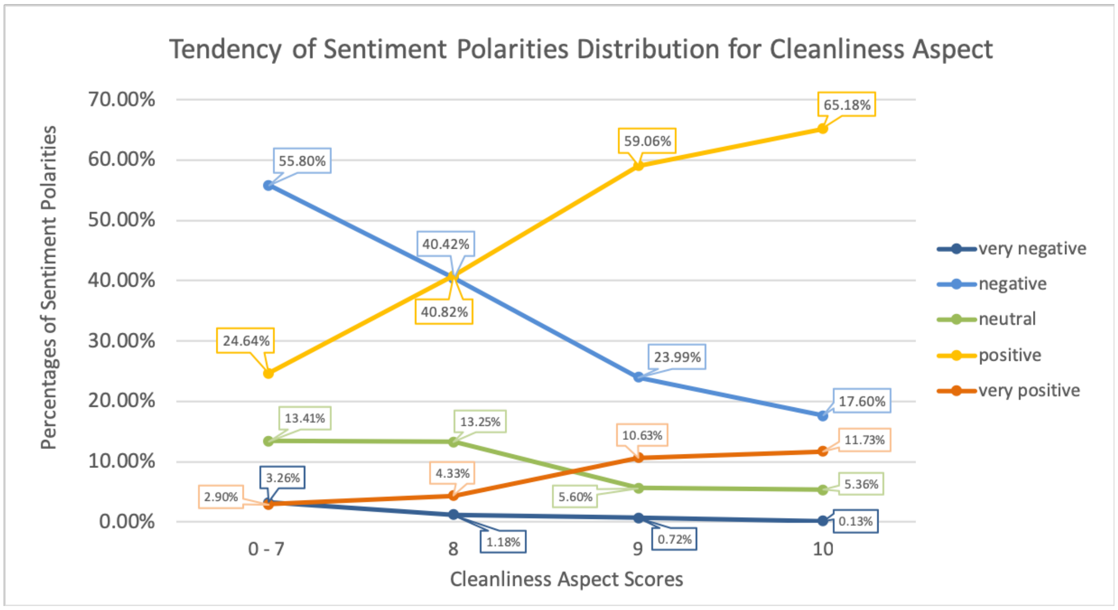

Figure 3: Tendency of Sentiment Polarities Distribution for Cleanliness Aspect

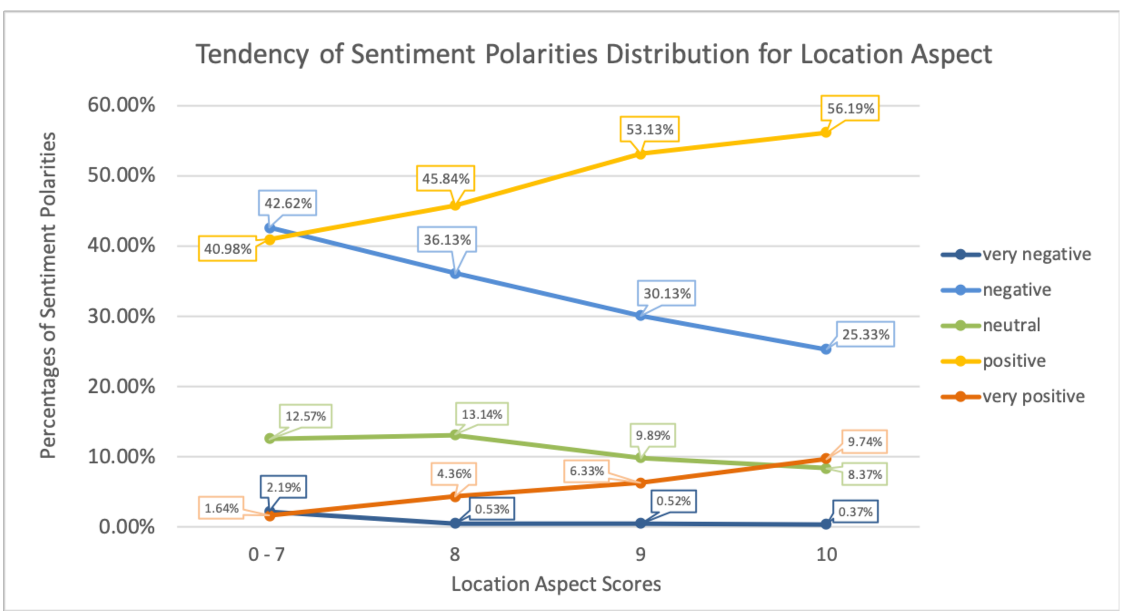

Figure 4: Tendency of Sentiment Polarities Distribution for Location Aspect

4.2.2 Average Sentiment Polarities Distribution of Aspects

The results presented in the previous section are based on the overall sentiment percentage of homes with different aspect score. However, it may be inaccurate because there is a possibility that the extreme situation such as one home having enough sentences with positive sentiment to overwhelm sentences with other polarities. To provide an alternate view, for every home, I computed the percentages of sentiment polarities for all aspect-relevant sentences, and then calculated the average of them for every score group.

ASPECT	POLARITIES	AVERAGE PERCENTAGES OF SENTENCES OF ALL HOMES			
		Score 0 – 7	Score 8	Score 9	Score 10
CLEANLINESS	very negative	5.16%	1.41%	0.55%	0.13%
	negative	50.35%	34.82%	21.48%	16.29%
	neutral	13.63%	7.58%	5.73%	5.63%
	positive	24.97%	49.35%	60.44%	66.99%
	very positive	5.89%	6.85%	11.80%	10.97%
LOCATION	very negative	1.05%	0.17%	0.69%	0.19%
	negative	41.34%	27.74%	30.66%	26.53%
	neutral	14.82%	11.90%	10.73%	8.14%
	positive	41.16%	52.09%	53.14%	55.95%
	very positive	1.64%	8.10%	4.77%	9.19%

Table 12: Average Percentages of Sentiment Polarities of Sentences of All Homes grouped by Aspect Scores

Table 12 and Figures 5 and 6 show the results and visualizations. We can see the overall trends (in Figure 5) for the cleanliness aspect are generally similar to the first analysis (in Figure 3). However, the proportion of very positive and neutral sentences increased slightly, while the very negative, correspondingly decrease for homes with lower aspect scores (0-7 or 8).

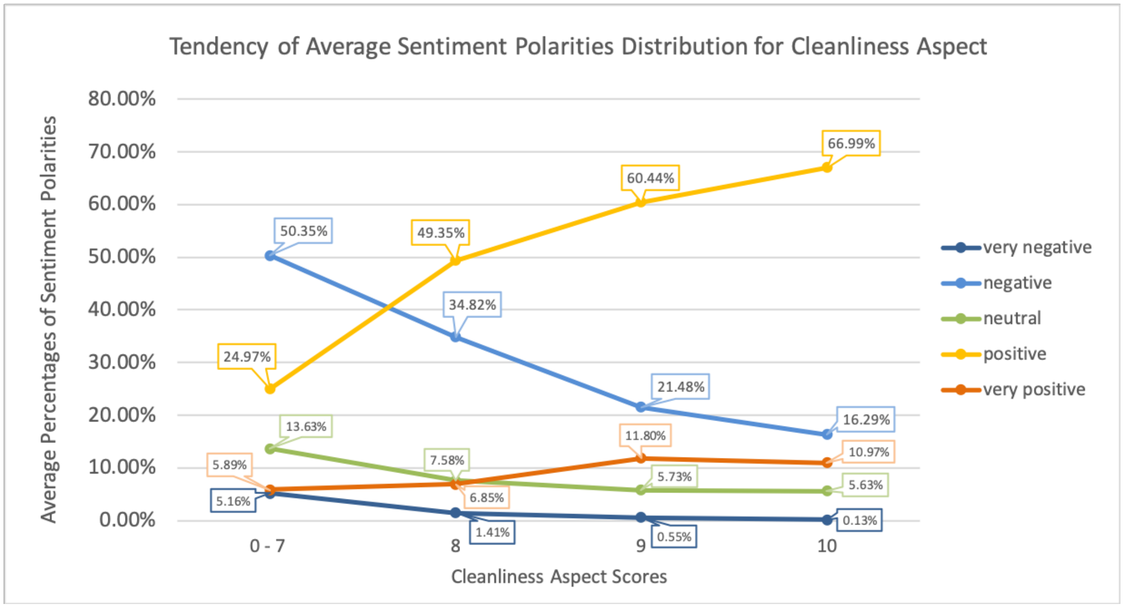

Figure 5: Tendency of Average Sentiment Polarities Distribution for Cleanliness Aspect

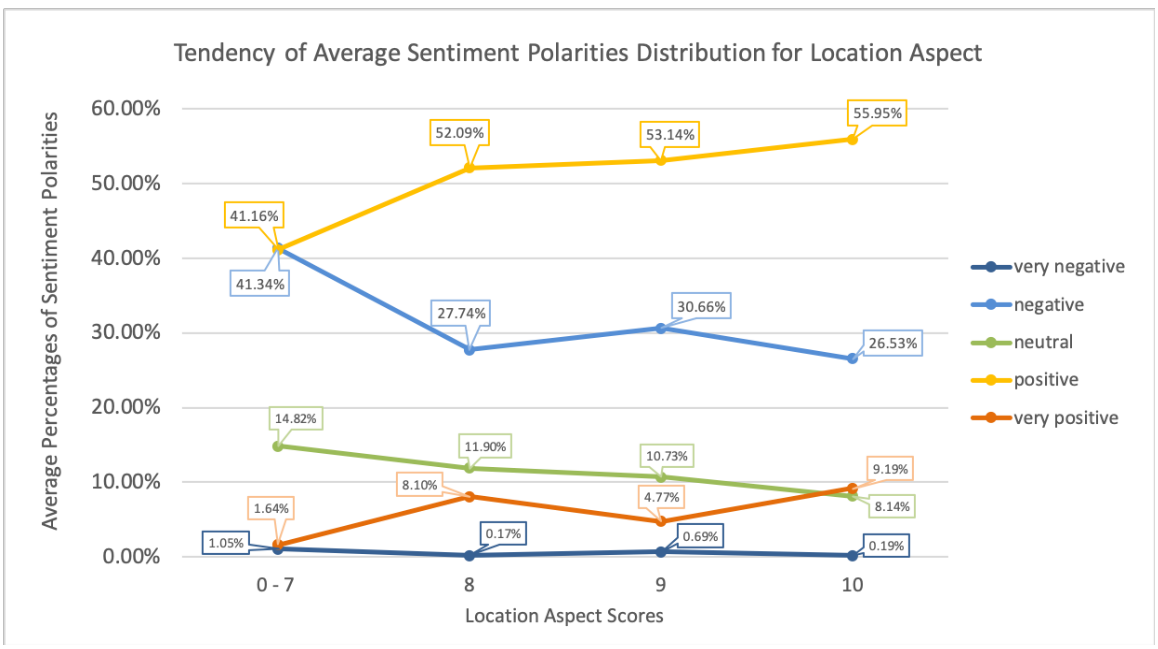

Figure 6: Tendency of Average Sentiment Polarities Distribution for Location Aspect

Things are a little different in the location aspect. As shown in Figure 6, the average percentages of all sentiment polarities basically keep stable for aspect scores 8, 9 and 10. The percentage of fluctuations of all sentiment polarities are all less than 5% (3.86%, 4.13%, 3.76%, 4.42%, 0.5% for positive, negative, neutral, very positive and very negative respectively). However, the polarities' overall percentages tendency (see Figure 4) shows larger fluctuations with 10.35%, 10.80%, 4.77%, 5.38%, 0.16% for positive, negative, neutral, very positive and very negative respectively. This means overall percentages of polarities enlarged the correspondence between aspect rating scores and the polarities' percentages. Homes with higher location scores did have larger proportion of positive sentences in reviews relevant to location aspect, but not as dramatically as in Figure 3. The trends reflected by average percentages (in Figure 6) should be more accurate.

4.2.3 Polarities Distribution Patterns for Homes with Same Aspect Score

Figures in this section (Figure 7, 8, 9, 10) display the data in Appendix I by plotting the sentiment polarities distribution of all relevant sentences of homes with the same aspect score. The figures for both aspects, cleanliness (Figure 7, 8) and location (Figure 9, 10), emphasize the difference between homes with scores of 0-7, and those with scores of 10. The area of red and yellow color representing very positive and positive in score 10 figures is obviously lager than it in score 0-7 figures. This appearance supports the discussion in previous section that homes with higher aspect score will have larger proportion of positive sentences and smaller proportion of negative sentences.

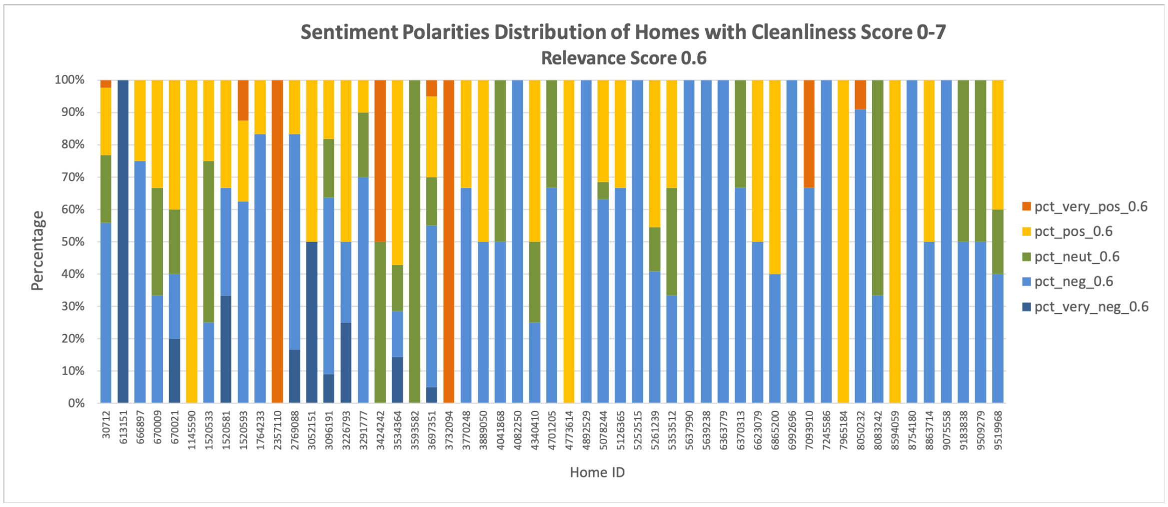

Figure 7: Sentiment Polarities Distribution of Homes with Cleanliness Score 0-7

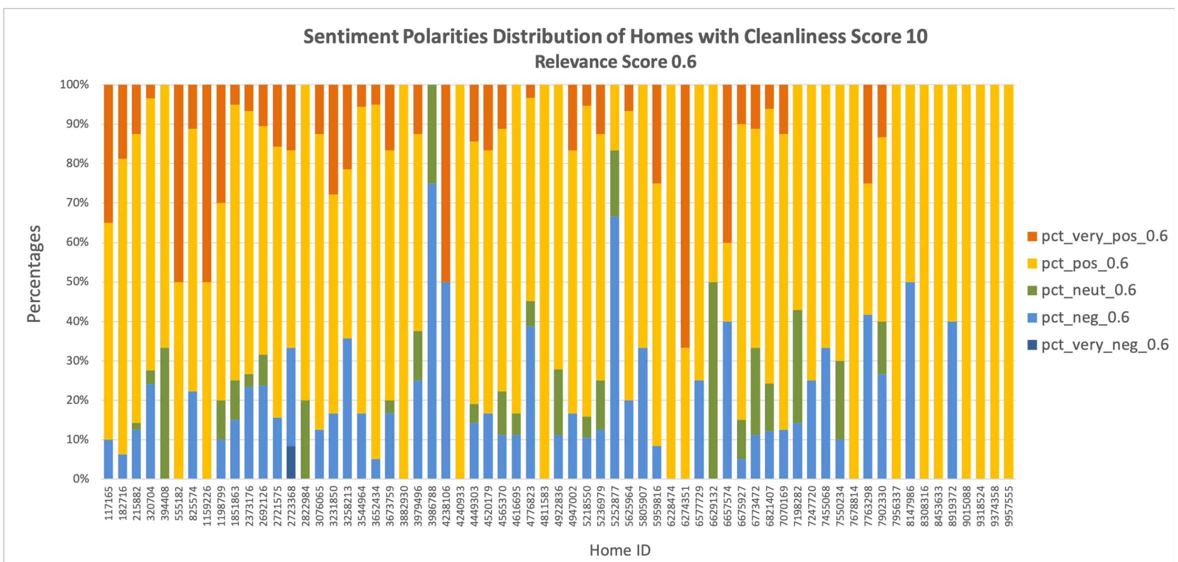

Figure 8: Sentiment Polarities Distribution of Homes with Cleanliness Score 10

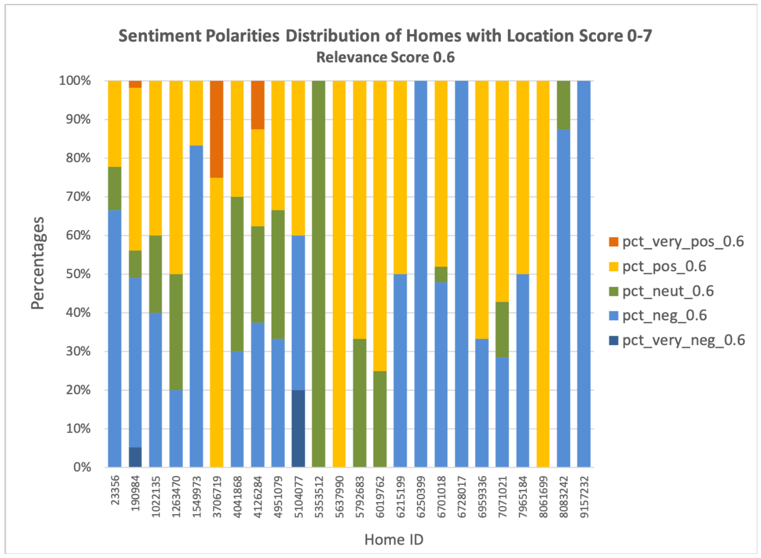

Figure 9: Sentiment Polarities Distribution of Homes with Location Score 0-7

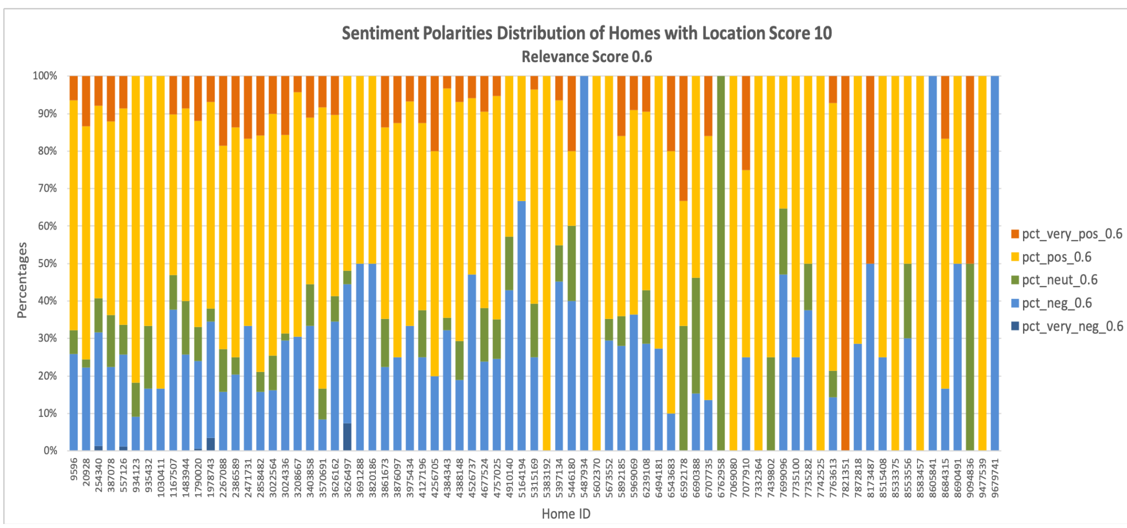

Figure 10: Sentiment Polarities Distribution of Homes with Location Score 10

5 CONCLUSION

This research analyzed the correspondence between review text and the scores of two aspects (cleanliness and location) to provide evidences based on sentence-level sentiment analysis. The results show that there exists a positive correlation between aspect scores and the proportion of positive aspect-relevant sentences in home reviews for both the aspects. The correlation in cleanliness analysis is high while in location analysis, the correlation became insignificant among score 8, 9 and 10. An additional finding is that reviewers may tend to mention or write more about an aspect if they rate the aspect a higher score.

In the methodology part, this paper proposed a sentence-to-aspect relevance detection approach for the subjectivity classification step of sentence-level sentiment analysis. Based on the manual validation, the approach is efficient to distinguish aspect relevant sentences from review paragraphs. Previous studies used the set of sentiment words with high occurrences in the text as lexicon in the subjectivity classification step, while this research provides a new approach of creating lexicon by opinion-target/aspect words and their synonyms and antonyms in WordNet. That is, for example, if one wants to analyze the sentiment of sentences that represent the “cleanliness” aspect of Airbnb homes, the lexicon of sentence subjectivity classification will consist of the word “cleanliness” and its synonyms and antonyms.

Future work could be in three directions. One is to use the proposed sentence-level sentiment analysis method on the other 4 aspects given by Airbnb: communication, check-in, value and accuracy. By comparing the results of all 6 aspects, one could achieve the evidence of a possible research question, “Do people pay more attention to some aspects than other aspects of Airbnb homes?”. Another direction is to deploy and generalize the methodology of this research to other kinds of online review data, for example, reviews of sneakers which have some conventional aspects like fit, style, etc. The possible research question could be “do the aspect scores of sneakers correspond to the reviews?”. The last direction could be the optimization and validation of the sentiment model proposed in this research. We could test more relevance scores in the subjectivity classification model to find a more precise threshold. The creation of the aspect-relevant lexicon is also a possible way to improve the model.

There are some limitations of this research. One is that the data size is still insufficient, only around 7000 reviews of each aspect. Analyzing more data could be helpful in both recognizing the distribution pattern among homes with a certain aspect score and improving the performance of the proposed model. Another limitation is that the data of this research are only the Seattle data. The results may be different for Airbnb homes in other regions or countries.

6 REFERENCES

Abdulla, N. A., Ahmed, N. A., Shehab, M. A., & Al-Ayyoub, M. (2013, December). Arabic sentiment analysis: Lexicon-based and corpus-based. In 2013 IEEE Jordan conference on applied electrical engineering and computing technologies (AEECT) (pp. 1-6). IEEE.

Barbosa, L., & Feng, J. (2010, August). Robust sentiment detection on twitter from biased and noisy data. In Proceedings of the 23rd international conference on computational linguistics: posters (pp. 36-44). Association for Computational Linguistics.

Bermingham, A., & Smeaton, A. (2011). On using Twitter to monitor political sentiment and predict election results. In Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011) (pp. 2-10).

Blair-Goldensohn, Sasha, Kerry Hannan, Ryan McDonald, Tyler Neylon, George A. Reis, and Jeff Reynar. Building a sentiment summarizer for local service reviews. In Proceedings of WWW-2008 workshop on NLP in the information explosion era. 2008.

Bo Pang, Lillian Lee, Shivakumar Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques", the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP'2002), 2002.

Bridges, J., & Vásquez, C. (2018). If nearly all Airbnb reviews are positive, does that make them meaningless?. Current Issues in Tourism, 21(18), 2057-2075.

Chaovalit, P., & Zhou, L. (2005, January). Movie review mining: A comparison between supervised and unsupervised classification approaches. In Proceedings of the 38th annual Hawaii international conference on system sciences (pp. 112c-112c). IEEE.

Cheng, M., & Jin, X. (2019). What do Airbnb users care about? An analysis of online review comments. International Journal of Hospitality Management, 76, 58-70.

Ding, X., Liu, B., & Yu, P. S. (2008, February). A holistic lexicon-based approach to opinion mining. In Proceedings of the 2008 international conference on web search and data mining (pp. 231-240). ACM.

Dave, K., Lawrence, S., & Pennock, D. M. (2003, May). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of the 12th international conference on World Wide Web (pp. 519-528). ACM.

El-Halees, A. M. (2011). Arabic opinion mining using combined classification approach. Arabic opinion mining using combined classification approach.

Esuli, A., & Sebastiani, F. (2006, May). Sentiwordnet: A publicly available lexical resource for opinion mining. In LREC(Vol. 6, pp. 417-422).

Hassan, A., Qazvinian, V., & Radev, D. (2010, October). What's with the attitude?: identifying sentences with attitude in online discussions. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (pp. 1245-1255). Association for Computational Linguistics.

Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 168-177). ACM.

Jiang, L., Yu, M., Zhou, M., Liu, X., & Zhao, T. (2011, June). Target-dependent twitter sentiment classification. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1 (pp. 151-160). Association for Computational Linguistics.

Joshi, M., Das, D., Gimpel, K., & Smith, N. A. (2010, June). Movie reviews and revenues: An experiment in text regression. In Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics (pp. 293-296). Association for Computational Linguistics.

Kessler, J. S., & Nicolov, N. (2009, March). Targeting sentiment expressions through supervised ranking of linguistic configurations. In Third International AAAI Conference on Weblogs and Social Media.

Kim, S. M., & Hovy, E. (2004, August). Determining the sentiment of opinions. In Proceedings of the 20th international conference on Computational Linguistics (p. 1367). Association for Computational Linguistics.

Kobayashi, N.; Iida, R.; Inui, K.; and Matsumoto, Y. 2006. Opin- ion Mining on the Web by Extracting Subject-Attribute-Value Relations. In AAAI-CAAW.

Le, Q., & Mikolov, T. (2014, January). Distributed representations of sentences and documents. In International conference on machine learning (pp. 1188-1196).

Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.

Lu, B., Ott, M., Cardie, C., & Tsou, B. K. (2011, December). Multi-aspect sentiment analysis with topic models. In 2011 IEEE 11th international conference on data mining workshops (pp. 81-88). IEEE.

M., von Hoffen, M., Hagge, M., Betzing, J. H., & Chasin, F. (2017). Leveraging social media to gain insights into service delivery: a study on Airbnb. Information Systems and e-Business Management, 1-23.

Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In LREc (Vol. 10, No. 2010, pp. 1320-1326).

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1–2), 1-135.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), 267-307.

Turney, P. D. (2002, July). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th annual meeting on association for computational linguistics (pp. 417-424). Association for Computational Linguistics.

Wang, S., & Manning, C. D. (2012, July). Baselines and bigrams: Simple, good sentiment and topic classification. In Proceedings of the 50th annual meeting of the association for computational linguistics: Short papers-volume 2 (pp. 90-94). Association for Computational Linguistics.

Wiebe, J., & Riloff, E. (2005, February). Creating subjective and objective sentence classifiers from unannotated texts. In International conference on intelligent text processing and computational linguistics (pp. 486-497). Springer, Berlin, Heidelberg.

Yu, H., & Hatzivassiloglou, V. (2003, July). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In Proceedings of the 2003 conference on Empirical methods in natural language processing (pp. 129-136). Association for Computational Linguistics.

Zervas, G., Proserpio, D., & Byers, J. (2015). A first look at online reputation on Airbnb, where every stay is above average. Where Every Stay is Above Average (January 28, 2015).

Appendix I

The Sentiment Polarities of Aspect-Relevant Sentences of Airbnb Homes with Cleanliness Score 0-7

home_id	# relevant sentences	# very negative sentences	# negative sentences	# neutral sentences	# positive sentences	# very positive sentences
30712	43	0	24	9	9	1
299817	0	0	0	0	0	0
613151	1	1	0	0	0	0
639130	0	0	0	0	0	0
666897	8	0	6	0	2	0
670009	3	0	1	1	1	0
670021	5	1	1	1	2	0
716829	0	0	0	0	0	0
1145590	1	0	0	0	1	0
1520533	4	0	1	2	1	0
1520581	3	1	1	0	1	0
1520593	8	0	5	0	2	1
1764233	6	0	5	0	1	0
2357110	1	0	0	0	0	1
2769088	6	1	4	0	1	0
2856806	0	0	0	0	0	0
3052151	2	1	0	0	1	0
3096191	11	1	6	2	2	0
3226793	4	1	1	0	2	0
3291777	10	0	7	2	1	0
3424242	2	0	0	1	0	1
3534364	7	1	1	1	4	0
3593582	1	0	0	1	0	0
3697351	20	1	10	3	5	1
3732094	1	0	0	0	0	1
3766285	0	0	0	0	0	0
3770248	3	0	2	0	1	0
3888924	0	0	0	0	0	0
3889050	2	0	1	0	1	0
4041868	2	0	1	1	0	0
4082250	1	0	1	0	0	0
4340410	4	0	1	1	2	0
4550099	0	0	0	0	0	0
4701205	3	0	2	1	0	0
4773614	1	0	0	0	1	0
4892529	1	0	1	0	0	0
4951079	0	0	0	0	0	0
5078244	19	0	12	1	6	0

5126365	3	0	2	0	1	0
5252515	1	0	1	0	0	0
5261239	22	0	9	3	10	0
5353512	3	0	1	1	1	0
5479566	0	0	0	0	0	0
5637990	5	0	5	0	0	0
5639238	2	0	2	0	0	0
6120046	0	0	0	0	0	0
6250399	0	0	0	0	0	0
6363779	2	0	2	0	0	0
6370313	3	0	2	1	0	0
6623079	2	0	1	0	1	0
6717555	0	0	0	0	0	0
6864319	0	0	0	0	0	0
6865200	5	0	2	0	3	0
6958436	0	0	0	0	0	0
6959336	0	0	0	0	0	0
6992696	6	0	6	0	0	0
7093910	3	0	2	0	0	1
7203765	0	0	0	0	0	0
7245586	1	0	1	0	0	0
7732071	0	0	0	0	0	0
7844444	0	0	0	0	0	0
7934356	0	0	0	0	0	0
7965184	1	0	0	0	1	0
7975026	0	0	0	0	0	0
7985714	0	0	0	0	0	0
8050232	11	0	10	0	0	1
8067053	0	0	0	0	0	0
8083242	3	0	1	2	0	0
8227710	0	0	0	0	0	0
8555304	0	0	0	0	0	0
8594059	1	0	0	0	1	0
8754180	1	0	1	0	0	0
8863714	2	0	1	0	1	0
8922554	0	0	0	0	0	0
8934054	0	0	0	0	0	0
9075558	7	0	7	0	0	0
9151374	0	0	0	0	0	0
9183838	2	0	1	1	0	0
9509279	2	0	1	1	0	0
9519968	5	0	2	1	2	0

The Sentiment Polarities of Aspect-Relevant Sentences of Airbnb Homes with Cleanliness Score 8

home_id	# relevant sentences	# very negative sentences	# negative sentences	# neutral sentences	# positive sentences	# very positive sentences
15108	7	0	3	0	3	1
258571	103	0	36	12	52	3
264829	15	0	3	2	8	2
278192	5	0	0	1	4	0
293890	12	0	4	2	6	0
338091	28	1	13	1	11	2
606297	28	0	12	5	10	1
637710	24	0	11	6	7	0
719233	127	2	41	26	52	6
859467	12	0	7	1	4	0
1039766	38	0	12	4	21	1
1107845	12	1	4	1	5	1
1190571	2	0	0	0	2	0
1541705	13	0	8	3	2	0
1589461	5	0	3	0	2	0
1709737	4	1	1	0	2	0
1724849	5	0	1	0	2	2
1773803	48	1	25	8	13	1
1815304	16	0	6	1	6	3
1831338	7	0	2	1	4	0
1840671	3	0	1	0	1	1
1950446	64	0	32	7	23	2
2909809	0	0	0	0	0	0
3155785	1	0	1	0	0	0
3528627	13	0	9	1	3	0
3562617	8	0	5	0	3	0
3732103	5	0	2	2	1	0
3811872	4	0	1	0	2	1
4130112	0	0	0	0	0	0
4258515	10	0	4	0	5	1
4264056	28	0	17	3	8	0
4384000	0	0	0	0	0	0
4632923	1	0	1	0	0	0
4681885	11	0	1	5	4	1
4716486	0	0	0	0	0	0
4808173	1	0	0	0	1	0
5062445	0	0	0	0	0	0
5364609	2	0	0	0	1	1
5459895	0	0	0	0	0	0
5487653	0	0	0	0	0	0
5618094	1	0	0	0	1	0
5931372	15	0	8	3	4	0
6278216	7	2	2	0	3	0
6337492	5	0	3	0	2	0
6425537	4	0	1	0	3	0
6436772	0	0	0	0	0	0
6528192	3	0	1	0	2	0
6705584	0	0	0	0	0	0

6852288	5	0	1	0	3	1
7095802	12	0	6	1	5	0
7219541	1	0	0	0	1	0
7219838	9	0	5	0	4	0
7386675	1	0	0	0	1	0
7401671	1	0	0	0	0	1
7634011	12	1	6	2	2	1
7649837	8	0	4	1	3	0
7667990	1	0	0	0	1	0
7940358	0	0	0	0	0	0
7970663	0	0	0	0	0	0
7987846	0	0	0	0	0	0
7999692	1	0	0	0	1	0
8147215	0	0	0	0	0	0
8355276	7	0	3	2	2	0
8391954	1	0	1	0	0	0
8517235	1	0	0	0	1	0
8553556	2	0	1	0	1	0
8578490	0	0	0	0	0	0
8631419	0	0	0	0	0	0
8843162	0	0	0	0	0	0
9134196	0	0	0	0	0	0
9157232	0	0	0	0	0	0
9303530	1	0	0	0	1	0
9411935	2	0	0	0	2	0
9511777	0	0	0	0	0	0
9532861	0	0	0	0	0	0

The Sentiment Polarities of Aspect-Relevant Sentences of Airbnb Homes with Cleanliness Score 9

home_id	# relevant sentences	# very negative sentences	# negative sentences	# neutral sentences	# positive sentences	# very positive sentences
13068	7	0	4	1	1	1
19611	13	0	7	0	5	1
86185	6	1	1	0	3	1
132120	10	0	1	0	9	0
286712	29	1	10	0	17	1
442487	43	0	9	4	25	5
444221	30	0	12	0	16	2
458186	5	0	1	0	4	0
479653	22	0	3	1	16	2
486344	11	0	2	1	6	2
486829	14	0	1	1	11	1
573942	22	0	6	1	13	2
654734	15	1	5	0	8	1
670056	13	0	4	1	8	0
1145148	4	0	2	0	2	0
1416763	5	0	2	1	2	0
1432713	56	0	13	0	34	9
1483944	21	0	4	2	12	3
1521633	7	0	2	2	2	1
1783382	9	0	0	0	7	2
2016613	2	0	0	1	1	0
2039149	56	0	12	8	33	3
2134911	0	0	0	0	0	0
2418658	26	0	7	2	11	6
2520890	5	0	3	0	2	0
2898401	12	0	3	2	6	1
2966415	16	0	6	0	9	1
2986056	0	0	0	0	0	0
3040278	33	1	4	1	21	6
3124383	8	0	1	0	6	1
3250577	1	0	0	0	1	0
3386862	23	0	9	1	12	1
3689416	7	0	1	0	5	1
3803212	1	0	0	0	1	0
4061051	2	0	1	0	1	0
4111954	2	0	0	1	1	0
4138423	10	0	3	0	6	1
4231670	2	0	0	0	2	0
4242626	1	0	0	0	0	1
4258762	13	1	1	0	8	3
4388148	19	0	2	1	15	1
4825472	4	0	0	0	4	0
4854767	5	0	1	1	3	0
5123904	2	0	1	0	1	0
5297143	1	0	0	0	0	1
5310193	1	0	0	0	1	0

5660792	13	0	5	0	7	1
6226666	2	0	0	0	2	0
6249536	1	0	0	0	1	0
6412566	4	0	0	1	3	0
6557297	4	0	3	0	1	0
6629278	5	0	3	0	2	0
6655233	5	0	2	0	2	1
6716620	0	0	0	0	0	0
6780670	4	0	1	1	1	1
6882518	8	0	0	0	7	1
6994406	0	0	0	0	0	0
7149703	5	0	0	0	3	2
7349099	1	0	0	0	1	0
7430926	12	0	2	1	8	1
7596455	0	0	0	0	0	0
7809718	1	0	0	0	1	0
7821003	1	0	0	0	1	0
7828222	1	0	1	0	0	0
7843145	3	0	0	0	2	1
7934963	8	0	1	2	5	0
7952930	7	0	2	0	4	1
8053921	3	0	1	0	1	1
8155710	0	0	0	0	0	0
8221520	4	0	0	0	2	2
8409975	5	0	1	1	3	0
9016362	2	0	1	0	0	1
9034515	0	0	0	0	0	0
9389755	2	0	0	0	2	0
9863484	1	0	0	0	1	0

The Sentiment Polarities of Aspect-Relevant Sentences of Airbnb Homes with Cleanliness Score 10

home_id	# relevant sentences	# very negative sentences	# negative sentences	# neutral sentences	# positive sentences	# very positive sentences
117165	20	0	2	0	11	7
182716	16	0	1	0	12	3
215882	56	0	7	1	41	7
320704	29	0	7	1	20	1
394408	9	0	0	3	6	0
555182	2	0	0	0	1	1
825574	18	0	4	0	12	2
1159226	6	0	0	0	3	3
1198799	10	0	1	1	5	3
1851863	20	0	3	2	14	1
2373176	30	0	7	1	20	2
2489283	0	0	0	0	0	0
2692126	38	0	9	3	22	4
2721575	32	0	5	0	22	5
2723368	12	1	3	0	6	2
2727984	0	0	0	0	0	0
2822984	5	0	0	1	4	0
3076065	8	0	1	0	6	1
3231850	36	0	6	0	20	10
3258213	14	0	5	0	6	3
3544964	18	0	3	0	14	1
3652434	20	0	1	0	18	1
3673759	30	0	5	1	19	5
3882930	2	0	0	0	2	0
3979496	8	0	2	1	4	1
3986788	4	0	3	1	0	0
4238106	2	0	1	0	0	1
4240933	1	0	0	0	1	0
4449303	21	0	3	1	14	3
4520179	6	0	1	0	4	1
4565370	9	0	1	1	6	1
4616695	18	0	2	1	15	0
4776823	31	0	12	2	16	1
4811583	2	0	0	0	2	0
4872699	0	0	0	0	0	0
4922836	18	0	2	3	13	0
4947002	6	0	1	0	4	1
5031357	0	0	0	0	0	0
5218550	19	0	2	1	15	1
5236979	8	0	1	1	5	1
5252877	6	0	4	1	1	0
5625964	15	0	3	0	11	1
5651254	0	0	0	0	0	0
5805907	3	0	1	0	2	0
5959816	12	0	1	0	8	3
6228474	1	0	0	0	1	0
6274351	3	0	0	0	1	2
6577729	8	0	2	0	6	0

6629132	2	0	0	1	1	0
6657574	5	0	2	0	1	2
6675927	20	0	1	2	15	2
6773472	9	0	1	2	5	1
6821407	33	0	4	4	23	2
7070169	8	0	1	0	6	1
7198282	7	0	1	2	4	0
7247720	8	0	2	0	6	0
7455068	3	0	1	0	2	0
7550234	10	0	1	2	7	0
7635966	0	0	0	0	0	0
7678814	2	0	0	0	2	0
7699356	0	0	0	0	0	0
7763298	12	0	5	0	4	3
7902330	15	0	4	2	7	2
7956337	1	0	0	0	1	0
8086294	0	0	0	0	0	0
8147986	2	0	1	0	1	0
8308316	3	0	0	0	3	0
8453633	1	0	0	0	1	0
8608233	0	0	0	0	0	0
8919372	5	0	2	0	3	0
9015088	1	0	0	0	1	0
9318524	1	0	0	0	1	0
9374358	3	0	0	0	3	0
9727857	0	0	0	0	0	0
9957555	1	0	0	0	1	0

The Sentiment Polarities of Aspect-Relevant Sentences of Airbnb Homes with Location Score 0-7

home_id	# relevant sentences	# very negative sentences	# negative sentences	# neutral sentences	# positive sentences	# very positive sentences
23356	9	0	6	1	2	0
190984	57	3	25	4	24	1
613151	0	0	0	0	0	0
1022135	5	0	2	1	2	0
1263470	10	0	2	3	5	0
1549973	6	0	5	0	1	0
3706719	4	0	0	0	3	1
4041868	10	0	3	4	3	0
4126284	8	0	3	2	2	1
4951079	3	0	1	1	1	0
5104077	5	1	2	0	2	0
5353512	2	0	0	2	0	0
5637990	2	0	0	0	2	0
5792683	3	0	0	1	2	0
6019762	4	0	0	1	3	0
6215199	4	0	2	0	2	0
6250399	1	0	1	0	0	0
6701018	25	0	12	1	12	0
6728017	1	0	1	0	0	0
6959336	3	0	1	0	2	0
7071021	7	0	2	1	4	0
7415378	0	0	0	0	0	0
7839723	0	0	0	0	0	0
7965184	4	0	2	0	2	0
8061699	1	0	0	0	1	0
8083242	8	0	7	1	0	0
9157232	1	0	1	0	0	0
9183838	0	0	0	0	0	0

The Sentiment Polarities of Aspect-Relevant Sentences of Airbnb Homes with Location Score 8

home_id	# relevant sentences	# very negative sentences	# negative sentences	# neutral sentences	# positive sentences	# very positive sentences
264829	27	0	7	5	15	0
278192	6	0	1	1	3	1
286712	37	1	15	3	16	2
385438	3	0	1	0	1	1
571640	4	0	1	2	1	0
571651	0	0	0	0	0	0
606297	57	0	28	6	22	1
609610	19	0	8	2	9	0
611500	19	0	6	6	6	1
666897	22	1	6	1	12	2
670021	8	0	5	1	2	0
693956	0	0	0	0	0	0
696004	47	0	23	6	17	1
716829	0	0	0	0	0	0
823989	33	0	12	4	17	0
877203	13	0	6	2	5	0
1179538	0	0	0	0	0	0
1246809	48	2	20	10	14	2
1340668	24	0	12	2	8	2
1461971	4	0	1	1	2	0
1484651	20	0	10	3	7	0
1499596	4	0	2	2	0	0
1566487	90	0	29	17	40	4
1571230	12	0	4	2	5	1
1652107	48	0	16	9	21	2
1672979	13	0	4	2	6	1
1815472	68	1	18	10	38	1
1905473	4	0	0	2	2	0
2016613	12	0	5	1	6	0
2056276	0	0	0	0	0	0
2134911	3	0	2	0	0	1
2508065	9	0	1	0	7	1
2520890	25	0	11	1	11	2
2586350	56	0	24	9	21	2
2586642	17	0	9	2	4	2
2610187	3	0	1	1	1	0
2800448	18	0	6	3	8	1
2980762	18	0	6	0	11	1
2986056	0	0	0	0	0	0
3291777	10	0	5	1	4	0
3303857	19	0	3	4	11	1
3316219	81	0	24	12	40	5
3329962	26	0	14	1	10	1
3449059	100	1	31	8	54	6
3533112	7	0	4	2	1	0
3544964	23	0	10	2	11	0
3630581	4	0	2	1	1	0
3726391	0	0	0	0	0	0

3849918	7	0	3	0	4	0
3859882	4	0	1	0	3	0
3904056	8	0	2	0	5	1
3939683	1	0	0	0	1	0
3951768	41	0	18	7	15	1
4105081	2	0	0	0	2	0
4130112	0	0	0	0	0	0
4144767	4	0	1	1	2	0
4163851	18	0	7	3	7	1
4340410	10	0	2	1	7	0
4395578	0	0	0	0	0	0
4395654	0	0	0	0	0	0
4589654	13	0	3	2	5	3
4708075	11	0	5	3	2	1
4825073	3	0	0	0	2	1
5021969	46	1	20	4	17	4
5126077	1	0	0	0	0	1
5219336	5	0	2	0	3	0
5252543	4	0	0	2	2	0
5340242	1	0	1	0	0	0
5365612	3	0	0	0	3	0
5376433	15	0	6	2	7	0
5407311	1	0	0	0	1	0
5424448	2	0	1	0	1	0
5471427	8	0	3	0	5	0
5744931	6	0	2	0	4	0
5931372	7	0	5	1	1	0
6133684	3	0	1	0	2	0
6226666	1	0	0	0	1	0
6249164	12	0	1	0	11	0
6291829	5	0	0	1	3	1
6325283	8	0	1	1	6	0
6411986	5	0	1	2	2	0
6416765	36	0	16	5	14	1
6425652	2	0	1	0	1	0
6545246	1	0	0	0	1	0
6714817	1	0	0	0	1	0
6864319	1	0	0	0	0	1
6907671	6	0	1	1	4	0
7007598	4	0	1	0	3	0
7013085	44	1	23	1	19	0
7027507	0	0	0	0	0	0
7048843	14	0	2	3	9	0
7095802	11	0	2	4	5	0
7203765	1	0	0	0	1	0
7205443	8	0	3	2	3	0
7219541	1	0	0	0	1	0
7388899	12	0	5	0	7	0
7420488	5	0	1	1	2	1
7429207	0	0	0	0	0	0
7431247	0	0	0	0	0	0
7500000	6	0	2	0	4	0
7697340	2	0	0	0	2	0

7902068	7	0	1	0	6	0
7922663	0	0	0	0	0	0
7934356	2	0	1	0	1	0
8168876	2	0	1	0	1	0
8342968	0	0	0	0	0	0
8525825	0	0	0	0	0	0
8799588	1	0	0	0	0	1
8848854	0	0	0	0	0	0
8921924	0	0	0	0	0	0
8934054	0	0	0	0	0	0
8952253	2	0	0	0	2	0
8988178	1	0	0	0	1	0
9075558	4	0	4	0	0	0
9134196	0	0	0	0	0	0
9151374	1	0	0	0	1	0
9170870	2	0	0	0	2	0
9186256	1	0	0	0	0	1
9199982	5	0	2	0	3	0
9238818	1	0	0	0	1	0
9300972	1	0	0	0	1	0
9449062	3	0	0	2	1	0
9473312	1	0	0	0	1	0
9507115	1	0	0	1	0	0
9519968	1	0	0	0	0	1
9532861	1	0	0	0	1	0
9564093	6	0	2	2	1	1
9736940	0	0	0	0	0	0

The Sentiment Polarities of Aspect-Relevant Sentences of Airbnb Homes with Location Score 9

home_id	# relevant sentences	# very negative sentences	# negative sentences	# neutral sentences	# positive sentences	# very positive sentences
202174	10	0	2	1	6	1
215882	142	0	35	11	82	14
219532	45	1	13	6	23	2
749758	9	0	3	0	6	0
762029	125	0	38	7	71	9
788146	15	0	4	0	7	4
1029680	1	0	0	1	0	0
1133329	85	1	25	4	44	11
1171574	141	1	34	21	76	9
1351716	120	1	34	13	69	3
1455908	32	0	8	4	17	3
1494986	9	0	3	1	5	0
1556170	65	0	26	5	32	2
1594313	10	0	1	0	9	0
1715582	22	0	8	2	9	3
2071305	10	0	5	0	5	0
2298775	15	0	5	0	9	1
2303981	4	0	1	0	3	0
2486612	59	0	16	6	32	5
3061918	22	0	8	1	12	1
3115801	3	0	2	0	1	0
3155785	0	0	0	0	0	0
3258213	32	0	13	0	15	4
3263722	2	0	1	1	0	0
3404700	48	0	13	8	26	1
3613516	1	0	0	0	1	0
3773784	22	0	2	3	17	0
3812725	3	0	0	0	3	0
3821663	94	0	29	19	41	5
4016650	30	2	12	2	14	0
4038347	34	0	9	4	21	0
4257097	1	0	1	0	0	0
4270642	14	0	6	1	6	1
4565370	22	0	6	2	11	3
5195064	40	0	18	2	16	4
5325809	16	0	6	0	9	1
5364609	5	0	0	1	3	1
5559063	1	0	0	0	1	0
6004058	17	0	5	3	8	1
6103108	0	0	0	0	0	0
6133354	24	0	7	4	11	2
6166307	21	0	9	0	11	1
6278680	4	0	0	0	4	0
6349598	8	0	1	1	5	1
6403104	0	0	0	0	0	0
6423204	39	0	11	1	26	1
6436772	3	0	0	0	3	0
6438013	4	1	1	0	2	0

6562393	7	0	2	0	3	2
6629278	9	1	3	0	5	0
6765354	11	0	2	3	6	0
6776229	11	0	1	4	6	0
6780615	6	0	3	1	2	0
6783394	48	1	22	4	16	5
6793490	29	0	9	2	17	1
6795678	22	0	11	1	9	1
6888107	11	0	3	2	6	0
6913414	10	0	2	1	7	0
7182584	3	0	2	0	1	0
7226013	9	0	3	0	6	0
7245586	4	0	0	1	3	0
7247518	4	0	0	2	2	0
7256122	21	0	8	2	11	0
7401671	10	0	2	0	6	2
7459637	6	0	2	2	2	0
7459684	15	0	5	1	6	3
7638282	10	0	3	1	4	2
7809455	9	0	4	0	5	0
8150145	7	0	3	1	3	0
8253955	2	0	1	0	1	0
8340819	1	0	1	0	0	0
8409969	17	0	4	3	10	0
8409975	4	0	2	0	2	0
8501705	19	0	3	5	11	0
8902887	5	0	2	1	2	0

The Sentiment Polarities of Aspect-Relevant Sentences of Airbnb Homes with Location Score 10

home_id	# relevant sentences	# very negative sentences	# negative sentences	# neutral sentences	# positive sentences	# very positive sentences
9596	31	0	8	2	19	2
20928	45	0	10	1	28	6
254340	76	1	23	7	39	6
387078	58	0	13	8	30	7
557126	175	2	43	14	101	15
934123	11	0	1	1	9	0
935432	6	0	1	1	4	0
1030411	6	0	1	0	5	0
1167507	98	0	37	9	42	10
1483944	35	0	9	5	18	3
1790020	175	1	41	16	96	21
1978743	29	1	9	1	16	2
2267088	70	0	11	8	38	13
2386589	44	0	9	2	27	6
2471731	6	0	2	0	3	1
2858482	19	0	3	1	12	3
3022564	130	0	21	12	84	13
3024336	51	0	15	1	27	8
3208667	23	0	7	0	15	1
3403858	9	0	3	1	4	1
3570691	24	0	2	2	18	2
3626162	29	0	10	2	14	3
3626497	27	2	10	1	14	0
3691288	4	0	2	0	2	0
3820186	2	0	1	0	1	0
3861673	125	0	28	16	64	17
3876097	8	0	2	0	5	1
3975434	15	0	5	0	9	1
4127196	8	0	2	1	4	1
4256705	10	0	2	0	6	2
4384343	31	0	10	1	19	1
4388148	58	0	11	6	37	4
4526737	17	0	8	0	8	1
4677524	21	0	5	3	11	2
4757025	57	0	14	6	34	3
4910140	14	0	6	2	6	0
5164194	6	0	4	0	2	0
5315169	28	0	7	4	16	1
5383192	1	0	0	0	1	0
5397134	31	0	14	3	12	2
5446180	5	0	2	1	1	1
5487934	1	0	1	0	0	0
5602370	1	0	0	0	1	0
5673552	17	0	5	1	11	0
5892185	25	0	7	2	12	4
5969069	11	0	4	0	6	1
6239108	21	0	6	3	10	2
6494181	11	0	3	0	8	0

6543683	10	0	1	0	7	2
6592178	3	0	0	1	1	1
6690388	13	0	2	4	7	0
6707735	44	0	6	0	31	7
6762958	1	0	0	1	0	0
7069080	2	0	0	0	2	0
7077910	12	0	3	0	6	3
7332364	1	0	0	0	1	0
7439802	4	0	0	1	3	0
7699096	17	0	8	3	6	0
7735100	4	0	1	0	3	0
7735282	8	0	3	1	4	0
7742525	1	0	0	0	1	0
7763613	14	0	2	1	10	1
7821351	1	0	0	0	0	1
7872818	14	0	4	0	10	0
8173487	2	0	1	0	0	1
8515408	4	0	1	0	3	0
8533375	4	0	0	0	4	0
8553556	10	0	3	2	5	0
8583457	4	0	0	0	4	0
8605841	1	0	1	0	0	0
8684315	6	0	1	0	4	1
8690491	10	0	5	0	5	0
9094836	2	0	0	1	0	1
9477539	1	0	0	0	1	0
9679741	1	0	1	0	0	0